

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2001, 15:28:47 ; Search time 2667.65 Seconds

(without alignments)  
2626.615 Million cell updates/sec

Title: US-09-784-340-3\_COPY\_18322\_18774

Perfect score: 453  
Sequence: 1 gtaagtaactactgtctgtac.....tgtgtgttttccttcacg 453

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_ba1:\*  
2: gb\_ba2:\*  
3: gb\_ba3:\*  
4: gb\_in1:\*  
5: gb\_in2:\*  
6: gb\_in3:\*  
7: gb\_om:\*  
8: gb\_ov:\*  
9: gb\_pat1:\*  
10: gb\_pat2:\*  
11: gb\_ph:\*  
12: gb\_pl1:\*  
13: gb\_pl2:\*  
14: gb\_pl3:\*  
15: gb\_pl4:\*  
16: gb\_ba1:\*  
17: gb\_ba2:\*  
18: gb\_fun:\*  
19: gb\_fun:\*  
20: gb\_fun:\*  
21: gb\_fun:\*  
22: gb\_fun:\*  
23: gb\_fun:\*  
24: gb\_fun:\*  
25: gb\_fun:\*  
26: gb\_fun:\*  
27: gb\_fun:\*  
28: gb\_fun:\*  
29: gb\_fun:\*  
30: gb\_fun:\*  
31: gb\_fun:\*  
32: gb\_fun:\*  
33: gb\_fun:\*  
34: gb\_fun:\*  
35: gb\_fun:\*  
36: gb\_fun:\*  
37: gb\_fun:\*  
38: gb\_fun:\*  
39: gb\_fun:\*  
40: gb\_fun:\*  
41: gb\_fun:\*  
42: gb\_fun:\*  
43: gb\_fun:\*

44: gb\_ov:\*  
45: gb\_pat:\*  
46: gb\_ph:\*  
47: gb\_pl:\*  
48: gb\_om:\*  
49: gb\_ov:\*  
50: gb\_fun:\*  
51: gb\_fun:\*  
52: gb\_fun:\*  
53: gb\_fun:\*  
54: gb\_fun:\*  
55: gb\_fun:\*  
56: gb\_fun:\*  
57: gb\_fun:\*  
58: gb\_fun:\*  
59: gb\_fun:\*  
60: gb\_fun:\*  
61: gb\_fun:\*  
62: gb\_fun:\*  
63: gb\_fun:\*  
64: gb\_fun:\*  
65: gb\_fun:\*  
66: gb\_fun:\*  
67: gb\_fun:\*  
68: gb\_fun:\*  
69: gb\_fun:\*  
70: gb\_fun:\*  
71: gb\_fun:\*  
72: gb\_fun:\*  
73: gb\_fun:\*  
74: gb\_fun:\*  
75: gb\_fun:\*  
76: gb\_fun:\*  
77: gb\_fun:\*  
78: gb\_fun:\*  
79: gb\_fun:\*  
80: gb\_fun:\*  
81: gb\_fun:\*  
82: gb\_fun:\*  
83: gb\_fun:\*  
84: gb\_fun:\*  
85: gb\_fun:\*  
86: gb\_fun:\*  
87: gb\_fun:\*  
88: gb\_fun:\*  
89: gb\_fun:\*  
90: gb\_fun:\*  
91: gb\_fun:\*  
92: gb\_fun:\*  
93: gb\_fun:\*  
94: gb\_fun:\*  
95: gb\_fun:\*  
96: gb\_fun:\*  
97: gb\_fun:\*  
98: gb\_fun:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	451.4	99.6	212904	66	AC021146 Homo sapi
2	359	79.2	138665	72	AC055794 Homo sapi
3	289	63.8	402	54	G50188 SHGC-79709
4	48.2	10.6	894	53	AL403234 T3 end of
5	48	10.6	161380	62	AC011683 Homo sapi
6	48	10.6	163233	80	AL356974 Homo sapi
7	48	10.6	182162	69	AC025302 Homo sapi
8	47.6	10.5	177341	68	AC023596 Homo sapi

Accession	Organism	Gene	Protein	Length	Quality
AF031413	Homo sapiens	COSPLP		1564	100
AL359252	Homo sapiens	DNA		1564	100
AC016462	Homo sapiens	DNA		1564	100
AC013504	Homo sapiens	DNA		1564	100
AL353783	Homo sapiens	DNA		1564	100
AC011254	Homo sapiens	DNA		1564	100
AL425150	clone BA0			1564	100
AL550243	Homo sapiens	DNA		1564	100
AL391538	Homo sapiens	DNA		1564	100
AL390793	Human DNA			1564	100
AL369912	Homo sapiens	DNA		1564	100
AL354831	Human DNA			1564	100
AC019350	Homo sapiens	DNA		1564	100
AL450447	Homo sapiens	DNA		1564	100
AC018674	Homo sapiens	DNA		1564	100
AL163817	Arabidopsis			1564	100
AL391626	Homo sapiens	DNA		1564	100
AL359928	Homo sapiens	DNA		1564	100
166494	Sequence 14			1564	100
AC010941	Homo sapiens	DNA		1564	100
AC012530	Homo sapiens	DNA		1564	100
AF298624	Dicotyle			1564	100
AL357043	Homo sapiens	DNA		1564	100
AL158044	Homo sapiens	DNA		1564	100
AC013382	Homo sapiens	DNA		1564	100
U00695	Yeast mtoc			1564	100
L36887	Saccharomyc			1564	100
AC012582	Homo sapiens	DNA		1564	100
AC020781	Homo sapiens	DNA		1564	100
AC019179	Homo sapiens	DNA		1564	100
AP002460	Arabidopsis			1564	100
AC023071	Homo sapiens	DNA		1564	100
AL157971	Homo sapiens	DNA		1564	100
AC021467	Homo sapiens	DNA		1564	100
AX059517	Sequence			1564	100
AC007730	Arabidopsis			1564	100

## ALIGNMENTS

RESULT	1
AC021146	
LOCUS	AC021146
DEFINITION	AC021146 212904 bp DNA HTG 07-JUL-2000
ACCESSION	AC021146
VERSION	AC021146.4 GI:8568861
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE	1 (bases 1 to 212904)
JOURNAL	Waterston, R.H.
REFERENCE	The sequence of Homo sapiens clone
AUTHORS	Unpublished
TITLE	2 (bases 1 to 212904)
JOURNAL	Waterston, R.H.
COMMENT	Direct Submission Submitted (14-JAN-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis MO 63108, USA On Jun 16, 2000 this sequence version replaced gi:7344259.

**Source**

```

Sequencing vector: M13: 88%
Sequencing vector: pIasmid: 12%
Chemistry: Dye-terminator Big Dye, 12% of reads
Chemistry: Dye-terminator B10 Dye, 12% of reads
Assembly program: Phrap: version 0.990319
Consensus quality: 199729 bases at least Q40
Consensus quality: 203731 bases at least Q30
Consensus quality: 206340 bases at least Q20
Insert size: 213000; agarose-fp
Insert size: 210604; sum-of-contigs
Quality coverage: 3.75 in Q20 bases; agarose-fp
Quality coverage: 3.83 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 24 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1      1564: contig of 1564 bp in length
*      1565      1664: gap of unknown length
*      1665      3085: contig of 1421 bp in length
*      3086      3185: gap of unknown length
*      3186      5134: contig of 1949 bp in length
*      5135      5234: gap of unknown length
*      5235      8169: contig of 2935 bp in length
*      8170      8699: gap of unknown length
*      8270      10341: contig of 2072 bp in length
*      10342      10441: gap of unknown length
*      10442      13614: contig of 3173 bp in length
*      13615      13714: gap of unknown length
*      13715      17089: contig of 3375 bp in length
*      17090      20701: gap of unknown length
*      17190      20701: contig of 3512 bp in length
*      20702      25001: gap of unknown length
*      20802      25001: contig of 4200 bp in length
*      25002      25101: gap of unknown length
*      25102      29020: contig of 3919 bp in length
*      29021      29120: gap of unknown length
*      29121      33356: contig of 4236 bp in length
*      33357      33456: gap of unknown length
*      33457      36902: contig of 3446 bp in length
*      36903      37002: gap of unknown length
*      37003      42975: contig of 5973 bp in length
*      42976      43075: gap of unknown length
*      43076      48339: contig of 5264 bp in length
*      48340      48439: gap of unknown length
*      48440      56629: contig of 8090 bp in length
*      56630      56629: gap of unknown length
*      56630      65213: contig of 8584 bp in length
*      65214      65314: gap of unknown length
*      65314      74715: contig of 9402 bp in length
*      74716      74815: gap of unknown length
*      74816      88546: contig of 13731 bp in length
*      88547      88646: gap of unknown length
*      88647      103367: contig of 14721 bp in length
*      103368      103467: gap of unknown length
*      103468      118167: contig of 14700 bp in length
*      118168      118267: gap of unknown length
*      118268      132765: contig of 14498 bp in length
*      132766      132865: gap of unknown length
*      132866      154092: contig of 21227 bp in length
*      154093      154192: gap of unknown length
*      154193      182145: contig of 27952 bp in length
*      182145      182245: gap of unknown length
*      182245      212904: contig of 30660 bp in length.
*
* Location/Qualifiers
*     1..212904
*         /organism="Homo sapiens"
*         /db_xref="taxon:9606"
*         /chromosome="4"

```







```

* 25164 25263: gap of 100 bp
* 25264 35051: contig of 9788 bp in length
* 35052 35151: gap of 100 bp
* 35152 48817: contig of 13666 bp in length
* 48818 48917: gap of 100 bp
* 48918 72679: contig of 23762 bp in length
* 72680 72779: gap of 100 bp
* 72780 108923: contig of 36144 bp in length
* 108924 109023: gap of 100 bp
* 109024 161380: contig of 52357 bp in length.
Location/Qualifiers
1. 161380
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="RP11-14M11"
/clone_1ib="RPci-11 Human Male BAC"
1. 2294
misc_feature
/feature="assembly-fragment"
2395..5926
/feature="assembly-fragment"
6027..9587
/feature="assembly-fragment"
9688..13188
/feature="assembly-fragment"
clone_end:17
vector_side:left"
13289..18750
/feature="assembly-fragment"
18851..25163
/feature="assembly-fragment"
25264..35051
/feature="assembly-fragment"
35152..48817
/feature="assembly-fragment"
clone_end:SP6
vector_side:right"
48918..72679
/feature="assembly-fragment"
72780..108923
/feature="assembly-fragment"
109024..161380
/feature="assembly-fragment"

```

```

BASE COUNT 52458 a 27890 c 28371 g 51661 t 1000 others
ORIGIN

```

```

Query Match 10.6%; Score 48; DB 62; Length 161380;
Best local Similarity 50.9%; Pred. No. 2.3;
Matches 114; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

OY 96 gatttactccaaatttgaagtcacacactgttactggaactgtggaatt 155
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 141475 GTTTATTGTAGACAGTATGATGATGTTTGTTCATTCGCAATGTTTAAAT 141416
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 156 gtatgataagatgtaacattcttcacatgaataatagtttaagtttaacaactgac 215
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 141415 GTAGTGTATATACATTTAATTAATATAGTACTAGTATAGTTACATTCACATCAAT 141356
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 216 ttactaagctttattacacatttaattaccacatttggtaagaataataactcttca 275
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 141355 ATTTTGTCTATTGCTTCACATTTGATATATCTATCTATCTTCATTCATTCATCTTTA 141296
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 276 gtctctcactatatactgttaataactgttaaccacaatatt 319
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 141295 TTTTGTCTCAGTAGATGTTTATATATGCAATTTAAATTAAT 141252
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

RESULT 6
LOCUS AL356974 163233 bp DNA HNG 07-APR-2001
DEFINITION Homo sapiens chromosome 1 clone RP11-14M11, *** SEQUENCING IN
PROGRESS ***, 5 unordered pieces.
ACCESSION AL356974 AC011683

```

```

VERSION AL356974.4 GI:11229160
KEYWORDS HNG, HTGS_PHASE1, HTGS_DRAFT, HTGS_FU11TOP.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 163233)
JOURNAL Plumb, B.
Direct Submission
Submitted (06-APR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
Requests: clonerequest@sanger.ac.uk
On Nov 20, 2000 this sequence version replaced gi:9800983.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: bal14M11
----- Summary Statistics
Assembly program: XCAP4; version 4.5
Sequencing vector: M13; M77815; 4% of reads
Sequencing vector: plasmid; 108752; 50% of reads
Chemistry: Dye-terminator ET-amersham; 1% of reads
Dye-terminator Big Dye; 98% of reads
Consensus quality: 161377 bases at least Q40
Consensus quality: 162006 bases at least Q30
Insert size: 162833; sum-of-contigs
Insert size: 147081; 12.5% error; agarose-fp
Quality coverage: 7.46x in Q20 bases; sum-of-contigs Quality
coverage: 8.46x in Q20 bases; agarose-fp
-----

```

```

Draft Sequence Produced by Whitehead Institute/MIT Center for
Genome Research, 320 Charles Street,
Cambridge, MA 02141, USA
http://www.seq.wi.mit.edu.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 29607: contig of 29607 bp in length
* 29608 29707: gap of 100 bp
* 29708 37686: contig of 7979 bp in length
* 37687 37786: gap of 100 bp
* 37787 139946: contig of 102160 bp in length
* 139947 140046: gap of 100 bp
* 140047 145961: contig of 5915 bp in length
* 145962 146061: gap of 100 bp
* 146062 163233: contig of 17172 bp in length.
Location/Qualifiers
1. 163233
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/clone_lib="RP11-14M11"
/clone_1ib="RPci-11.1"
1. 29607
/feature="assembly-fragment:02718
fragment_chain:1
clone_end:SP6
vector_side:left"
29708..37686
/feature="assembly-fragment:00563
fragment_chain:1"
37787..139946
/feature="assembly-fragment:01001
fragment_chain:2"

```

```

FEATURES
source
misc_feature
misc_feature
misc_feature

```

misc_feature	140047..145961	/note="assembly_fragment:00462 fragment_chain:2"
misc_feature	146062..163233	/note="assembly_fragment:00430 fragment_chain:2"
BASE COUNT	54411 a 28101 c 29009 g 51305 t	407 others
ORIGIN		
Query Match	10.6%; Score 48; DB 80; Length 163233;	
Best Local Similarity	50.9%; Pred. No. 2.3;	
Matches 114; Conservative	0; Mismatches 110; Indels 0; Gaps 0;	
Oy	96 gatttactccaatattgtagcttcaatcattcaacctgttactggaatagtgtggaat	155
Db 114443	gTTTATTGTAGACACTAATGATCATGCTTTTGTTCATTCTGCCAATGTTGTTTAAT	114384
Oy	156 gtagctacagagtgctcaaccttcttcgtgaatattagttttaagttacacactgc	215
Db 114383	GTAGTGTTTAATACATTTAATTTAATATAGTACTGATAAGTTTATGATTCACATCTATT	114324
Oy	216 ttaccaagcttttaccatcattcaatttaccacatttgttgaagaatcattcttca	275
Db 114323	ATTTTGGCATTTGGCTTTCACACTTTGATATATCTTATCTATTTTCCATTACTCTTTA	114264
Oy	276 gtcttcacatactatcgtttaactatgttaccacaataat	319
Db 114263	TTTTGTGTCACTAGATGATTTTATATATATAGCCATTTAATTAATTACT	114220
RESULT 7		
AC025302	182162 bp	DNA
LOCUS	Homo sapiens clone RP11-680K19, WORKING DRAFT SEQUENCE, 14	HTG 25-MAY-2000
DEFINITION	unordered pieces.	
ACCESSION	AC025302	GI:8072561
VERSION	AC025302.3	GI:8072561
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
AUTHORS	1 (bases 1 to 182162)	
TITLE	Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beka,F., Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G., Campoliano,A., Castle,A., Choepey,J., Colangelo,M., Collins,S., Colliomore,A., Cooke,P., Dearlano,K., Dewar,K., Diaz,J.S., Dodge,S., Domingo,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gaidyna,S., Glnde,S., Goyette,M., Graham,L., Grand-Pierre,N., Grant,G., Hagos,B., Heathford,A., Horton,L., Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Labocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K., MacDonald,P., Marquis,N., McCarthy,M., McEwan,P., McGurt,A., McKernan,K., McPheters,R., Melidiri,T., Meneses,L., Mihova,T., Miranda,C., Mleaga,V., Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N., Pisanl,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Seery,P., Spencer,B., Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Testafae,S., Theodore,J., Tittel,I., Travers,M., Triggillo,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and zodj,M.	
JOURNAL	Unpublished	
REFERENCE	2 (bases 1 to 182162)	
AUTHORS	Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beka,F., Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G., Campoliano,A., Castle,A., Choepey,J., Colangelo,M., Collins,S., Colliomore,A., Cooke,P., Dearlano,K., Dewar,K., Diaz,J.S., Dodge,S., Domingo,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gaidyna,S., Glnde,S., Goyette,M., Graham,L., Grand-Pierre,N., Grant,G., Hagos,B., Heathford,A., Horton,L., Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Labocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K., MacDonald,P., Marquis,N., McCarthy,M., McEwan,P., McGurt,A., McKernan,K., McPheters,R., Melidiri,T., Meneses,L., Mihova,T., Miranda,C., Mleaga,V., Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N., Pisanl,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Seery,P., Spencer,B., Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Testafae,S., Theodore,J., Tittel,I., Travers,M., Triggillo,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and zodj,M.	
TITLE	Direct Submission	
JOURNAL	Submitted (08-MAR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA	
COMMENT	On May 25, 2000 this sequence version replaced g1:7280312.	

```

All repeats were identified using RepeatMasker:
Smith, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L4330
Center clone name: 680_K.19
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 176346 bases at least Q40
Consensus quality: 179048 bases at least Q30
Consensus quality: 180129 bases at least Q20
Insert size: 183000; agarose-fp
Insert size: 180862; sum-of-contigs
Quality coverage: 5.0 in Q20 bases; agarose-fp
Quality coverage: 5.1 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of 'N', but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1.1322: contig of 1322 bp in length
* 1323 1422: gap of 100 bp
* 1423 4198: contig of 2776 bp in length
* 4199 4298: gap of 100 bp
* 4299 6627: contig of 2329 bp in length
* 6628 6727: gap of 100 bp
* 6728 9902: contig of 3175 bp in length
* 9903 10002: gap of 100 bp
* 10003 13206: contig of 3204 bp in length
* 13207 13306: gap of 100 bp
* 13307 17106: contig of 3800 bp in length
* 17107 17206: gap of 100 bp
* 17207 21975: contig of 4769 bp in length
* 21976 22075: gap of 100 bp
* 22076 24992: contig of 2917 bp in length
* 24993 25092: gap of 100 bp
* 25093 33458: contig of 8346 bp in length
* 33459 33538: gap of 100 bp
* 33539 49399: contig of 15661 bp in length
* 49400 49499: gap of 100 bp
* 49500 69413: contig of 19914 bp in length
* 69414 69513: gap of 100 bp
* 69514 91950: contig of 22437 bp in length
* 91951 92050: gap of 100 bp
* 92051 119326: contig of 27276 bp in length
* 119327 119426: gap of 100 bp
* 119427 182162: contig of 62736 bp in length.
Location/Qualifiers
1.182162
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="RP11-680K19"
/clone_lib="RP11-11 Human Male BAC"
1.1322
/note="assembly_fragment"
1423.4198
/note="assembly_fragment"
4299.6627
/note="assembly_fragment"
6728.9902
/note="assembly_fragment"
10003.13206

```

```

misc_feature /note="assembly-fragment"
13307..17106
/note="assembly-fragment"
17207..21975
/note="assembly-fragment"
22076..24992
/note="assembly-fragment"
clone_end:17
vector_side:left"
25093..33438
/note="assembly-fragment"
33539..49399
/note="assembly-fragment"
clone_end:SP6
vector_side:left"
49500..69413
/note="assembly-fragment"
69514..91950
/note="assembly-fragment"
92051..119326
/note="assembly-fragment"
119427..182162
/note="assembly-fragment"
BASE COUNT 58607 a 31962 c 32284 g 58008 t 1301 others
ORIGIN
Query Match 10.6% Score 48; DB 69; Length 182162;
Best Local Similarity .50.9%; Pred. No. 2.2;
Matches 114; Conservative 0; Mismatches 110; Indels 0; Gaps 0;
QY 96 gatattaccacacatattgagtcattcattacacattgcttcaggaatattggaatt 155
DB 16544 GTTATTGTGACAGTAAGATGATGATGTTTGTTCATTCTGCCAATGTTTAAAT 16603
QY 156 gtaattcatagagtcacactcttcacatggaatattgtaagttaacacactgac 215
DB 16604 GTAGTCTTAAATCAATTTAAATTAATAGTACTGATGATGATGATGATGATGAT 16663
QY 216 ttactaagctttattacacattacattacacacatttctgtaagaatactcttca 275
DB 16664 ATTTGCTATTCCTTTCACACATTTGATATATCTATCTATCTATCTATCTATCT 16723
QY 276 gtcttcacatattctgtttactactgttaacacacattt 319
DB 16724 TTTTGTTCAGTAGATGTTTATATATATGCCATTAAATTAC 16767
RESULT 8
AC023596 177341 bp DNA HTG 04-NOV-2000
LOCUS Homo sapiens chromosome 3 clone RP11-699J21, WORKING DRAFT
DEFINITION SEQUENCE, 17 unordered pieces.
ACCESSION AC023596
VERSION AC023596.17 GI:11079253
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 177341)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,
Alshrocks,S.L., Amaralunga,H.C., Are,D.R., Banks,T., Barbarella,J.,
Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowie,S., Brieve,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burke,P., Burkett,C., Burrell,K.L., Byrd,N.C., Caron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhury,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,

```

# TITLE JOURNAL REFERENCE AUTHORS TITLE JOURNAL COMMENT

```

* Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T.,
* Garza, N., Gill, R., Gorrell, J.H., Guevara, M., Gunaratne, P., Hale, S.,
* Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A.,
* Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C.,
* Hollins, B., Homs, F., Howard, S., Huber, J., Huliy, S., Hume, J.,
* Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S.,
* Joudan, C., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J.,
* Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, J.C.,
* Lewis, L., Li, J., Li, Z., Licharge, O., Lien, C., Liu, J., Liu, W.,
* Louisa, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R.,
* Ma, J., Maheshwari, M., Mapa, P., Martin, R., Martindale, A.,
* Martinez, E., Massey, E., Mawhney, E., McLeod, M.P., Meador, M.,
* Mei, G., Metzger, M., Miner, G., Miner, Z., Mitchell, T., Mohabadi, K.,
* Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N.,
* Nguyen, A., Nguyen, S., Nguyen, N., Nickerson, E., Nwokoko, S.,
* Ogih, M., Okunoda, G., Oragunye, N., Oyedele, R., Pace, A., Payton, B.,
* Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Py, L.L.,
* Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojudo, A., Rolfe, M.,
* Ruiz, S., Saverio, G., Scherer, S., Scott, G., Shen, H., Shooshbari, N.,
* Sisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H.,
* Stone, H., Sutton, A., Swalek, A., Taber, P., Tamerisa, A., Tamerisa, K.,
* Tang, H., Tansey, J., Taylor, C., Taylor, V., Villalona, D., Vinson, R.,
* Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalona, D., Vinson, R.,
* Wall, R., Wang, S., Ward, Moore, S., Warren, R., Wasylyuk, R., Wooden, S.,
* Watlington, S., Williams, G., Williamson, G., Wleczek, R., Woodson, D.,
* Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
* and Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 177341)
Worley, K.C.
Direct Submission
Submitted (16-FEB-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 3, 2000 this sequence version replaced gi:10304992.
----- Genome Center of Medicine
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: RP11-699J21
Center clone name: RP11-699J21
Sequencing vector: M13; L08821
Chemistry: Dye-terminator Big Dye; 50% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 166534 bases at least Q40
Consensus quality: 169833 bases at least Q30
Estimated insert size: 171835; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 4.2x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 32477: contig of 32477 bp in length
* 32478 32577: gap of unknown length
* 32578 51994: contig of 19417 bp in length
* 51995 52094: gap of unknown length
* 52095 71484: contig of 19390 bp in length
* 71485 71584: gap of unknown length
* 71585 84749: contig of 13165 bp in length

```





```

RES
source
Location/Qualifiers
1. 139104
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="RP11-377K21"
/clone_lib="RPC1-11 Human Male BAC"
1. 237
/misc_feature
/note="assembly_fragment"
clone_end:17
vector_side:left"
338. 1699
/misc_feature
/note="assembly_fragment"
1800. 2964
/misc_feature
/note="assembly_fragment"
3065. 4681
/misc_feature
/note="assembly_fragment"
4782. 6187
/misc_feature
/note="assembly_fragment"
6288. 8839
/misc_feature
/note="assembly_fragment"
8940. 11261
/misc_feature
/note="assembly_fragment"
11362. 13371
/misc_feature
/note="assembly_fragment"
13472. 15817
/misc_feature
/note="assembly_fragment"
15918. 18313
/misc_feature
/note="assembly_fragment"
18414. 20632
/misc_feature
/note="assembly_fragment"
20733. 23241
/misc_feature
/note="assembly_fragment"

```

Query Match	10.4%	Score 47.2	DB 67	Length 139104
Best Local Similarity	48.2%	Pred. No. 3.3		
Matches 133	Conservative	0	Mismatches 143	Indels 0
QY 169	tgcaaaccttcttcacatggaatathaggtttagcgaacgaactggtcctaagcttt	228		
Db 11206	TATATTAATTTATTAATGCTAAATATTTAAATTTAAATATTAATTAATTAATTAATTA	11447		
QY 229	attcacactcttaatttcaccactttgttgaagaatacacttcacgctctccactat	288		
Db 11146	TTACATATTTGTATACATTATATTTATTAATAATATATATTTATATATTTATATTTATAT	11087		
QY 289	atctgtttaactatgtaaccacaatatctcactgcaacaccgaactcaactctttat	348		
Db 11086	ATTAATTTATTTTATTAATAAAACATATTAATTTATATTAATTAATTAATTAATTAATTA	11027		
QY 349	gaacagtcttcgctgcgtcgaatacactacggtttctacagctgctttatgaa	408		
Db 11026	TATATATTTATTAATTAATAATATATATATATATTAATTTATATATTAATTAATTAATTT	10967		
QY 409	caaaactacaactcttcaagctctatgctgctgtttt	444		
Db 10966	ATATATTAATTTATTAATTAATTTATTAATTAATGATTTT	10931		

RESULT 10			
AF031433	1355 bp	DNA	PLN
LOCUS			
DEFINITION	Gossypium arboreum tRNA-Thr (trnT (UGU)) and tRNA-Ieu (trnL (UAA)) partial genes, chloroplast genes encoding chloroplast RNA, partial sequence, and trnT (UGU)-trnL (UAA) intergenic region.		
ACCESSION	AF031433		

VERSION	AF031433.1	GI:2623897
KEYWORDS	Gossypium arboreum:	
SOURCE	Chloroplast Gossypium arboreum	
ORGANISM	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta	
REFERENCE	Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II	
AUTHORS	Malvales; Malvaceae; Gossypium.	
TITLE	1 (bases 1 to 1335)	
JOURNAL	Small,R.L., Ryburn,J.A., Cronn,R.C., Seelanan,T. and Wendel,J.F.	
REFERENCES	The tortoise and the hare: Choosing between noncoding plastome and nuclear adh sequences for phylogeny reconstruction in a recently diverged plant group	
AUTHORS	Am. J. Bot. 85 (9), 1301-1315 (1998)	
TITLE	2 (bases 1 to 1335)	
JOURNAL	Small,R.L., Ryburn,J.A., Cronn,R.C., Seelanan,T. and Wendel,J.F.	
REFERENCES	Directed Submission	
TITLE	Submitted (27-OCT-1997) Botany, Iowa State University, 353 Bessey Hall, Ames, IA 50011, USA	
FEATURES	Location/Qualifiers	
Source	1..1335	
misc_feature	/organism="Gossypium arboreum"	
gene	/organelle="plastid;chloroplast"	
tRNA	/db_xref="taxon:29729"	
misc_feature	complement(<1..20)	
gene	/gene="trnI (UGU)"	
tRNA	/product="tRNA-Thr"	
misc_feature	complement(<1..20)	
gene	/gene="trnT (UGU)"	
tRNA	21..1335	
misc_feature	/note="intergenic spacer"	
gene	1335..>1355	
tRNA	/gene="trnL (CAA)"	
gene	/product="tRNA-Leu"	
BASE COUNT	1354..>1355	
ORIGIN	554 a 139 c 175 g 487 t	
Query Match	10.4%; Score 47; DB 12; Length 1355;	
Best Local Similarity	48.3%; Pred. No. 4.8;	
Matches	131; Conservative 0; Mismatches 140; Indels 0; Gaps 0;	
OY	107 acaatattgagtcattcacccttgtaactggaatagtgtggaatatgtgatcatag 166	
Db	185 ATAACTTAATATAATATAACATATTATTAATATAATAGTAGAATCAACTGAATATTC 244	
OY	167 agtgcacaactttcttcatacgaaataattaaggtttaagttaacaaactgctctaagctt 226	
Db	245 AAGCTAATGTGATAGAATCTTAATATTCGATTTAAATTTAATATATATATATTTA 304	
OY	227 ttattcacactcaatttatcccccatttgttaagaataatactcttcagctctcccaat 286	
Db	305 TTATTTAATCTATATATTTATATATTTAATATATATTTAATATATCTTTTTTTATATCTAGTT 364	
OY	287 atatctgttlaactaactaactaaccaaatattcatactgtaacacagaatcaatcttta 346	
Db	365 ATATTTCTTTTATATATTAATATAGTAATATTTACTATADATATACGACGCCATTTTATTTTC 424	
OY	347 ctgaacatgctcttgctgctgataacata 377	
Db	425 TAGTTTATTTTATATTTATTTAGTATATTTTACA 455	
RESULT	11	
LOCUS	AL359262/c	
DEFINITION	Human DNA sequence from clone RP11-334O13 on chromosome 13,	
ACCESSION	complete sequence.	
VERSION	AL359262	
KEYWORDS	AL359262.9 GI:12964305	
SOURCE	HIG.	
SOURCE	human	

ORGANISM	Homo sapiens
REFERENCE	Mammalia; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;
AUTHORS	Enxayivola; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE	1 (bases 1 to 162763)
JOURNAL	Tracey/A.
COMMENT	Direct Submission Submitted (01-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Requests: clonerequests@sanger.ac.uk On Feb 19, 2001 this sequence version replaced g1:12718063. During sequence assembly data is computed from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality > 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw: SWISSPROT; Tr:, TREMBL; Wp:, WORPEP; Information on the WORPEP database can be found at: <a href="http://www.sanger.ac.uk/Projects/C_elegans/worpep/">http://www.sanger.ac.uk/Projects/C_elegans/worpep/</a> This sequence was generated from part of bacterial clone contigs of human chromosome 13, constructed by the Sanger Centre Chromosome 13 Mapping Group. Further information can be found at <a href="http://www.sanger.ac.uk/MGP/Chr13">http://www.sanger.ac.uk/MGP/Chr13</a> RP11-334013 is from the library RPci-11.2 constructed by the group of Pieter de Jong. For further details see <a href="http://www.chori.org/bacpac/home.htm">http://www.chori.org/bacpac/home.htm</a> VECTOR: pBAC3.6 This sequence is the entire insert of clone RP11-334013 The true left end of clone RP11-105A24 is at 125885 in this sequence. The true right end of clone RP11-52F22 is at 10371 in this sequence.
FEATURES	Location/unlifiers
SOURCE	1..162763
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="13"
	/clone="RP11-334013"
	/clone_lib="RPCI-11.2"
	134..171
repeat_1_region	/note="MER66B repeat: matches 405..442 of consensus"
	348..508
repeat_2_region	/note="LTR43 repeat: matches 132..296 of consensus"
	660..1735
repeat_3_region	/note="T1 repeat: matches 3870..4955 of consensus"
	1741..2133
repeat_4_region	/note="MT2CB repeat: matches 1..423 of consensus"
	2139..2812
repeat_5_region	/note="L1MA8 repeat: matches 5158..5850 of consensus"
	2814..3014
repeat_6_region	/note="L1MB2 repeat: matches 5952..6163 of consensus"
	3526..3684
repeat_7_region	/note="MER5A repeat: matches 30..189 of consensus"
	3758..3916
repeat_8_region	/note="LTR1D repeat: matches 352..499 of consensus"
	3917..4221
repeat_9_region	/note="AlusX repeat: matches 1..307 of consensus"
	4222..4330
repeat_10_region	/note="LTR1D repeat: matches 198..352 of consensus"
	4436..4487
repeat_11_region	/note="8 copies 4 mer acac 90% conserved"
	7714..7944
repeat_12_region	/note="Mir repeat: matches 2..249 of consensus"
	8322..8866
repeat_13_region	/note="MER41A repeat: matches 3..554 of consensus"
	10653..10836

```
repeat_region /note="92 copies 2 mer tt 59% conserved"
10808. 10863
repeat_region /note="14 copies 4 mer cctc 71% conserved"
11705. 12067
repeat_region /note="MER47A repeat: matches 3. 366 of consensus"
13870. 14231
repeat_region /note="MER67D repeat: matches 1. 340 of consensus"
18418. 18688
repeat_region /note="Aluub repeat: matches 5. 283 of consensus"
22171. 22452
repeat_region /note="L1MC3 repeat: matches 7429. 7735 of consensus"
23271. 23298
repeat_region /note="7 copies 4 mer aaca 92% conserved"
24848. 25683
repeat_region /note="L1 repeat: matches 2350. 3164 of consensus"
25785. 26264
repeat_region /note="240 copies 2 mer ta 55% conserved"
26320
misc_feature /note="Weak data"
26490. 26633
repeat_region /note="72 copies 2 mer ta 60% conserved"
26511. 26634
repeat_region /note="31 copies 4 mer atat 61% conserved"
26634. 26679
repeat_region /note="23 copies 2 mer at 76% conserved"
26796. 26847
repeat_region /note="26 copies 2 mer ga 76% conserved"
27188. 27379
repeat_region /note="L1M4 repeat: matches 4614. 4812 of consensus"
27730. 28333
repeat_region /note="L1M3A repeat: matches 5298. 5951 of consensus"
30264. 30320
repeat_region /note="L2 repeat: matches 2606. 2661 of consensus"
30627. 31083
repeat_region /note="L2 repeat: matches 2020. 2527 of consensus"
31378. 31662
repeat_region /note="Alusg repeat: matches 23. 307 of consensus"
32773. 32812
repeat_region /note="L2 repeat: matches 2238. 2277 of consensus"
33003. 33130
repeat_region /note="L2 repeat: matches 2538. 2674 of consensus"
33237. 33784
repeat_region /note="MLT2A repeat: matches 1. 453 of consensus"
34490. 34696
repeat_region /note="L1M1 repeat: matches 5407. 5613 of consensus"
34698. 34925
repeat_region /note="L1 repeat: matches 5168. 5403 of consensus"
34925. 35042
repeat_region /note="L1 repeat: matches 4513. 4630 of consensus"
35045. 35531
repeat_region /note="L1 repeat: matches 3187. 3672 of consensus"
35531. 37715
repeat_region /note="L1M2 repeat: matches -673. 1545 of consensus"
43406. 43672
repeat_region /note="Alusg1 repeat: matches 1. 296 of consensus"
43830. 44106
repeat_region /note="L1M2 repeat: matches 16. 304 of consensus"
44438. 44717
repeat_region /note="L2 repeat: matches 2405. 2703 of consensus"
44762. 45209
repeat_region /note="L2 repeat: matches 1813. 2289 of consensus"
45279. 45583
repeat_region /note="L2 repeat: matches 1229. 1554 of consensus"
45639. 45797
repeat_region /note="MIR repeat: matches 81. 248 of consensus"
47207. 48663
repeat_region /note="L1M1 repeat: matches 4500. 6117 of consensus"
48811. 48902
repeat_region /note="MIR repeat: matches 70. 163 of consensus"
49688. 49982
repeat_region /note="Alusg repeat: matches 2. 302 of consensus"
50030. 50123
repeat_region /note="47 copies 2 mer tc 76% conserved"

repeat_region 50039. 50118
/note="20 copies 4 mer cctc 82% conserved"
repeat_region 52618. 52916
/note="L2 repeat: matches 179. 504 of consensus"
repeat_region 52975. 53182
/note="L2 repeat: matches 645. 858 of consensus"
repeat_region 53309. 53359
/note="MER53 repeat: matches 139. 189 of consensus"
repeat_region 53423. 53515
/note="MER53 repeat: matches 29. 122 of consensus"
repeat_region 53517. 53634
/note="59 copies 2 mer tt 61% conserved"
repeat_region 53714. 54065
/note="L2 repeat: matches 1172. 1561 of consensus"
repeat_region 55215. 55388
/note="MER20 repeat: matches 1. 163 of consensus"
repeat_region 56456. 56696
/note="Aluub repeat: matches 1. 235 of consensus"
repeat_region 57571. 57918
/note="L2 repeat: matches 1996. 2380 of consensus"
repeat_region 58095. 58201
/note="L2 repeat: matches 2634. 2750 of consensus"
repeat_region 58275. 58312
/note="19 copies 2 mer tt 78% conserved"
repeat_region 58353. 58412
/note="15 copies 4 mer gaaa 75% conserved"
repeat_region 59013. 59610
/note="MER4B repeat: matches 1. 575 of consensus"
repeat_region 60115. 60597
/note="MLT1D repeat: matches 3. 501 of consensus"
repeat_region 60600. 60946
/note="L2 repeat: matches 525. 873 of consensus"
repeat_region 61270. 61299
/note="15 copies 2 mer aa 93% conserved"
repeat_region 61349. 61723
/note="MLT1B repeat: matches 8. 390 of consensus"
repeat_region 63420. 63548
/note="MIR repeat: matches 119. 254 of consensus"
repeat_region 65053. 65557
/note="MER84 repeat: matches 2. 508 of consensus"
repeat_region 65558. 65629
/note="MER84 repeat: matches 1. 74 of consensus"
repeat_region 66347. 66822
/note="MER52A repeat: matches 545. 746 of consensus"
repeat_region 66868. 67056
/note="LTR28 repeat: matches 1. 174 of consensus"
repeat_region 67112. 67208

Query Match 10.1%; Score 45.8; DB 90; Length 162763;
Best Local Similarity 47.9%; Pred. No. 6.1;
Matches 195; Conservative 0; Mismatches 207; Indels 5; Gaps 2;
```



OY	78	ctcgtagacatglttgaggagtttactccacaatttgagtcattcaacacctgttact	137
Db	165719	TATATTATAATATTAAATATTAATACAAATATTATATATATATATATATAT	165778
OY	138	ggaatagttgtggaaatttgttaagtcataagaigtgc aaacttctcatggaaatltag	197
Db	165779	ACAATATTATTAAATTTATTTAATAATATTTTAAATATTAATACAAATATTAAAA	165838
OY	198	ttaagttaaacaactggctcactagaagtttatcacactcaatctaacccatttgt	257
Db	165839	TATATTATATATTTTATTCACAAATATTATTAATATTAATACCAATATTAATATAT	165898
OY	258	taagaatatactcttcacgtctctccactatc-----tgttaactctgttaaccac	313
Db	165899	ATATATATTATTAAATTTATTAATACAAATATTAAATATATATTTATTAATATATATAC	165958
OY	314	aatactatgccaacacgaatacaatcttctactg-aaatgltcttggtctgacaaac	372
Db	165959	ATTATTAATACCAATATATTAATATATATTTTAAATACAAATATTATATATTTATAAT	166018
OY	373	atactacacggttctcactagtgcttttagaaaacaacactacaa	419
Db	166019	AATATATCAAAATATTTTAAATAGTATTAATATATATATTCACAAATATTTAA	166065

RESULT	13
AC013504	
LOCUS	AC013504
DEFINITION	Homo sapiens clone RP11-115G14, WORKING DRAFT SEQUENCE.
ACCESSION	AC013504
VERSION	AC013504.3
KEYWORDS	HTGS; HTGS_PHASE1; HTGS_DRAFT.
SOURCE	human.
ORGANISM	Homo sapiens

REFERENCE	1 (bases 1 to 186695)	Nusbaum, C., and Lander, E.
AUTHORS	Birren, B., Linton, L.,	
TITLE	Homo sapiens, clone RP11-115G14	
JOURNAL	Unpublished	
REFERENCE	2 (bases 1 to 186695)	
AUTHORS	Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M., Brown, J., Braithwaite, R., Brockway, S., Canham, C., Cook, D., Doolittle, R., Finkbeiner, S., Gillanders, A., Gillis, J., Glavins, T., Hendrickson, E., Jones, R., Kagan, J., Kohly, J., Kreitzer, A., LaBaere, J., Landrum, P., Lau, Y., Leach, M., Liebman, I., Lin, X., Little, L., Liu, X., Luo, M., Ma, Z., McEwen, H., Miller, M., Morgan, B., Mullikin, J., Nelson, D., Nishizawa, N., Okazaki, Y., Orntoft, T., Park, S., Paterson, G., Peterson, J., Phillips, R., Reid, J., Richmond, C., Samadpour, M., Schaeffer, N., Scheet, O., Shao, C., Smith, D., Sprague, A., Stange, C., Steinberg, M., Stewart, J., Sulikowski, J., Taniguchi, T., Tompa, R., White, P., Wierzbicki, A., Wu, X., Yan, H., Yao, J., Zhang, J., Zheng, X., and Ziegler, W.	

TITLE Direct Submission  
JOURNAL Submitted (13-NOV-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
COMMENT On Sep 9, 2000 this sequence version replaced g1:6573941.  
3. Replaced with Bioethelmark:

Genomic  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: <http://www-seq.wi.mit.edu>  
 Contact: [sequence.submissions@genome.wi.mit.edu](mailto:sequence.submissions@genome.wi.mit.edu)  
 ----- Project Information -----  
 Center project name: I1865  
 Center clone name: I15\_G-14  
 ----- Summary Statistics -----

## FEATURES

**Source**

```

Sequencing vector: M13; M7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 154341 bases at least Q40
Consensus quality: 173898 bases at least Q30
Consensus quality: 180811 bases at least Q20
Insert size: 167000; agarose-fp
Insert size: 185495; sum-of-ctrlsigs
Quality coverage: 4.7 in Q20 bases; agarose-fp
Quality coverage: 4.2 in Q20 bases; sum-of-ctrlsigs

```

---

```

NOTE: This is a 'working draft' sequence. It currently
consists of 13 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

```

```

1      26180: contig of 26180 bp in length
      26181 26280: gap of 100 bp
      26281 27621: contig of 1341 bp in length
      27622 27721: gap of 100 bp
      27722 29672: contig of 1951 bp in length
      29673 29772: gap of 100 bp
      29773 34488: contig of 4616 bp in length
      34489 34488: gap of 100 bp
      34489 41138: contig of 6350 bp in length
      41139 41538: gap of 100 bp
      41539 57120: contig of 15582 bp in length
      57121 57220: gap of 100 bp
      57221 75067: contig of 17847 bp in length
      75068 75167: gap of 100 bp
      75168 93217: contig of 18050 bp in length
      93218 93317: gap of 100 bp
      93318 114597: contig of 21280 bp in length
      114598 114697: gap of 100 bp
      114698 133497: contig of 18800 bp in length
      133498 133597: gap of 100 bp
      133598 154392: contig of 20795 bp in length
      154393 154492: gap of 100 bp
      154493 178008: contig of 23576 bp in length
      178009 178168: gap of 100 bp
      178169 186695: contig of 8527 bp in length.

```

```
misc_feature  
l. 26180 /note="assembly-fragment  
clone.end:sp6  
vector.side:left"  
26281..27621  
misc_feature /note="assembly-fragment"  
27722..29672  
/note="assembly-fragment"  
29773..34388  
/note="assembly-fragment"  
34489..41438  
/note="assembly-fragment"  
41539..57120  
misc_feature /note="assembly-fragment"  
57221..75067  
/note="assembly-fragment"  
75168..93217  
/note="assembly-fragment"  
93318..114537  
/note="assembly-fragment"  
114698..133497  
misc_feature /note="assembly-fragment"  
133598..153492  
misc_feature
```

```

misc_feature      /note="assembly_fragment"  

                  154493..178068  

                  /note="assembly_fragment"  

misc_feature      178169..186695  

                  /note="assembly_fragment  

                   clone_end:r7  

                   vector_side:right"  

BASE COUNT       58821 a 32954 c 32688 g 61023 t 1209 others  

ORIGIN  

Query Match      10.1%; Score 45.8; DB 63; Length 186695;  

Best Local Similarity 47.9%; Pred. No. 6.1;  

Matches 195; Conservative 0; Mismatches 207; Indels 5; Gaps 2  

OY 18 tacagacgtgacttaacaattgactaigtatatacatatataccagaanaagttaaatacacc 77  

    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||  

Db 25256 TATATAAATACAAATATATTATTAATATAATATAATATAATATAATATAATATAATATA 25315  

OY 78 ctggcagacagctgttgagggaatttacctccacaaatatgtgcattcataccctgttact 137  

    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||  

Db 25316 TATTATATATATTTTATTAATATATTAATATACAAATATATTATATATATATATATATAT 25375  

OY 138 ggaatcagctgtcggaattgttagtcatacagagtcgaacctttccatcggaatatatagg 197  

    |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  

Db 25376 ACAATATATTTAATAATATATTATATATATATATATATATATATATATATATATATATATAT 25435  

OY 198 tttaagcttaacaacgctgcttactaaacctllttaaccaatlttaaccacctttgt 257  

    |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  

Db 25436 TATTATATATATATATATATATATCAAAATATATTTATTAATATATATATATATATATATAT 25495  

OY 258 taagaataactctcttcagctctcccaataac-----tgttaatacacatgaaccaac 313  

    ||||||||| ||||||||| ||||||||| ||||||||| ||||||||| ||||||||| ||  

Db 25496 ATATATATATATATATATATATATATATACAAATATTTTATTAATATATATATATATATAT 25555  

OY 314 aaatcatgctcacaaacccaalcaactctttactg-aacatgctcttgctgcataac 372  

    |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  

Db 25556 ATTATTAATACAAATATATTTATTAATATATATATATTAATTAATACAAATATATTAATTAAT 25615  

OY 373 atatactacggttlctacacgfgctctllttaaaaacaactaacaa 419  

    |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  

Db 25616 ATATATATCAAAATATTTATTAAGTGTAATATATTAATATACAAATATATATA 25662  

RESULT 14  

AL353783 182232 bp DNA HTG: 06-MAR-2001  

LOCUS Homo sapiens chromosome 9 clone RP11-271O3, *** SEQUENCING IN  

DEFINITION PROGRESS ***, 5 unordered pieces.  

ACCESSION AL353783  

VERSION AL353783.6 GI:13273652  

KEYWORDS HTG; HUGS_PHASE1; HUGS_DRAFT.  

SOURCE human.  

ORGANISM Homo sapiens  

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  

REFERENCE 1 (bases 1 to 182232)  

AUTHORS Plumb,B.  

TITLE Direct Submissions  

JOURNAL Submitted (01-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire,  

CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  

REQUESTS: Clonerquest@sanger.ac.uk  

On Mar 11, 2001 this sequence version replaced gi:13234872.  

COMMENT ----- Genome Center  

Center: Sanger Centre  

Center code: SC  

Web site: http://www.sanger.ac.uk  

Contact: humquery@sanger.ac.uk  

Project Information  

Center project name: ba27103  

Summary Statistics  

Assembly program: XGAP4, version 4.5  

Sequencing vector: plasmid;L08752; 100% of reads
```

```

Chemistry: Dye-terminator Big Dye: 100% of reads
Consensus quality: 180394 bases at least Q40
Consensus quality: 181011 bases at least Q30
Consensus quality: 181436 bases at least Q20
Insert size: 181832; sum-of-contigs
Insert size: 173385; 3.0% error; agarose-fp
Quality coverage: 8.57x in Q20 bases; sum-of-contigs
Quality coverage: 9.10x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 47401: contig of 47401 bp in length
* 47402 47501: gap of 100 bp
* 47502 82489: contig of 34988 bp in length
* 82490 82589: gap of 100 bp
* 82590 150067: contig of 67478 bp in length
* 150068 150167: gap of 100 bp
* 150168 155571: contig of 5404 bp in length
* 155572 155671: gap of 100 bp
* 155672 182232: contig of 26561 bp in length.
Location/Qualifiers
1..182232
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/clone="RP11-27103"
/clone_id="RPC1-11.1"
1..47401
/note="assembly_fragment:00307
clone_end:SP6
vector_side:left"
47502..82489
/note="assembly_fragment:00972
fragment_chain:1"
82590..150067
/note="assembly_fragment:03273
fragment_chain:1"
150168..155571
/note="assembly_fragment:02310"
155672..182232
/note="assembly_fragment:02328"
BASE COUNT 59904 a 33351 c 30550 g 58027 t 400 others
ORIGIN
Query Match 10.1%; Score 45.6; DB 79; Length 182232;
Best Local Similarity 47.8%; Pred. No. 6.6;
Matches 132; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
QY 169 tgcacaacttcttcataatgaagaatataaggttaagcctaaacacgtgctactaagccttt 228
Dp 156077 TATATAAATATATAAATATATAAATATATAAATATATAAATATATAAATATATAAATATA 156136
QY 229 attcacacttaatttaccacatttgcttaagaatactcttcagctctccactact 288
Dp 156137 TTACATATTTGTATGCATTTATATATATATATAAATATATAATATATATATATATATAT 156196
QY 289 atctgttaataactatgtaaacacaataatcatgctcacccacaagaatcaacttact 348
Dp 156197 ATTAATTTATTTATATATAAATATATAATTTATATATAAATATAATTAATATATATAATTA 156256
QY 349 gaacatgctcttgctgcatcaacaatactactaagcttatactcagctgctcttatacgaaa 408
Dp 156257 TATATAATTATATATAATATATAATATATAATTTAAATATATATAATTAATATATAATTT 156316
QY 409 caaactacacttcttaagttctatgctgttttt 444

```





us-09-784-340-3\_copy\_18322\_18774.rge

Page 17

Search completed: August 27, 2001, 17:57:00  
Job time: 8893 sec

**THIS PAGE BLANK (USPTO)**